Cell PRESS
Structure
Article

Fragile X Mental Retardation Syndrome: Structure of the KH1-KH2 Domains of Fragile X Mental Retardation Protein

Roberto Valverde,[1] Irina Pozdnyakova,[1] Tommi Kajander,[1,3] Janani Venkatraman,[1] and Lynne Regan[1,2,*]
[1]Department of Molecular Biophysics and Biochemistry
[2]Department of Chemistry
Yale University, 266 Whitney Avenue, New Haven, CT 06520, USA
[3]Present address: Macromolecular X-Ray Crystallography, Research Program in Structural Biology and Biophysics, Institute of Biotechnology, P.O. Box 65, University of Helsinki, 00014 Helsinki, Finland.
*Correspondence: lynne.regan@yale.edu
DOI 10.1016/j.str.2007.06.022

SUMMARY

Fragile X syndrome is the most common form of inherited mental retardation in humans, with an estimated prevalence of about 1 in 4000 males. Although several observations indicate that the absence of functional Fragile X Mental Retardation Protein (FMRP) is the underlying basis of Fragile X syndrome, the structure and function of FMRP are currently unknown. Here, we present an X-ray crystal structure of the tandem KH domains of human FMRP, which reveals the relative orientation of the KH1 and KH2 domains and the location of residue Ile304, whose mutation to Asn is associated with a particularly severe incidence of Fragile X syndrome. We show that the Ile304Asn mutation both perturbs the structure and destabilizes the protein.

INTRODUCTION

Fragile X Mental Retardation Protein (FMRP) is expressed at high levels in the central nervous system and genitalia (Tamanini et al., 1997). Intracellularly, FMRP localizes predominantly in the cytoplasm (Devys et al., 1993), where it has been reported to associate with transcribing ribosomes; low levels of FMRP are also detectable in the nucleolus (Jin and Warren, 2000). The sequence of FMRP provides hints to its function. FMRP contains nuclear localization and nuclear export signals (NLS and NES) (Eberhart et al., 1996), tandem K-homology domains (KH1 and KH2) (Siomi et al., 1993), and an RGG box (Figure 1A). However, despite significant experimental efforts, the in vivo activity of FMRP remains unclear, with roles in nuclear-cytoplasmic shuttling of RNA, translational control, dendritic transport of RNA, and dendrite-specific regulation of translation all having been proposed (Jin and Warren, 2000; Eberhart et al., 1996; Khandjian et al., 2004; Antar et al., 2005).

Humans and other mammals have two autosomal paralogs of FMRP: Fragile X Related Proteins 1 and 2 (FXRP1 and FXRP2), which have similar domain organization and expression patterns to FMRP. Their functions are also unknown. Animals in other phyla either have no homologs of FMRP or have a single FMRP-like ortholog, as seen, for example, in amphibians and arthropods (Figure 1B).

The most common cause of Fragile X syndrome is an expansion of CGG repeats upstream of the gene that encodes FMRP, and the consequent silencing of gene expression essentially eliminates production of FMRP (Figure 1A). There are, however, examples of individuals who do not have the upstream expansion, but rather have deletions or point mutations within the FMRP gene (Hammond et al., 1997). An individual was identified with an especially severe manifestation of Fragile X syndrome—low IQ, severe social and behavioral impairment, macroorchidism—who had the single point mutation (Ile304Asn) within the KH2 domain of FMRP (De Boulle et al., 1993). These observations indicate a key function for FMRP in normal cells and highlight the importance of the KH domains. There have been a number of suggestions as to the possible effects of the Ile304Asn mutation on FMRP function, but no definitive conclusion (Siomi et al., 1994; Zhang et al., 2001; Wan et al., 2000).

KH modules are widespread, versatile nucleic acid-binding motifs that most often occur in tandem arrays of between 2 and 16 repeats. Sequence alignments show that, typically, the first KH domain (KH1) in a protein is more similar to other KH1 domains in different proteins than to KH2, KH3, etc. domains within the same protein. The same result is found for KH2 and KH3 domains. The minimal KH motif is comprised of 45 amino acids and is characterized by a $\beta1\alpha1\alpha2\beta2$ fold, with a consensus *GXXG* loop connecting the two central helices ($\alpha1$ and $\alpha2$). KH domains are classified as Type I or Type II folds. Both contain the minimal KH motif, but with different C- or N-terminal extensions (underlined) giving $\beta1\alpha1\alpha2\beta1\underline{\beta'\alpha'}$ and $\underline{\alpha'\beta'}\beta1\alpha1\alpha2\beta2$ for Type I and Type II, respectively (Figure 2). KH domains in eukaryotic proteins are exclusively Type I, whereas those in prokaryotic proteins are exclusively Type II (Siomi et al., 1993; Grishin, 2001). Although

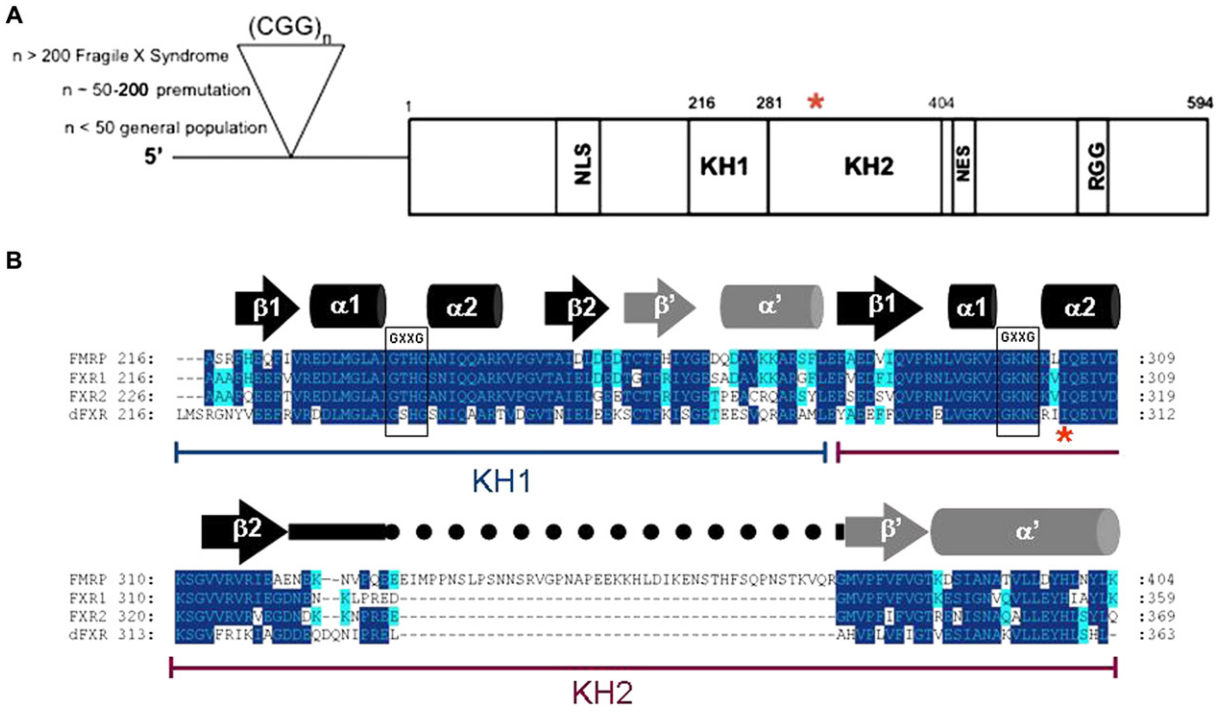

Figure 1. Fragile X Mental Retardation Protein

(A) Domain organization of human FMRP. The red asterisk indicates the location of the point mutation (Ile304Asn), which results in a severe Fragile X Mental Retardation phenotype in humans.

(B) Sequence alignment of the KH1 and KH2 domains of human FMRP, its autosomal paralogs FXRP1 and FXRP2, and the *Drosophila* ortholog, dFXRP. Residues that are identical in three or more of the proteins are colored dark blue; residues that are identical in two of the proteins are colored light blue. The numbering pertains to the full-length proteins. The KH1 and KH2 domains of hFMRP are underscored in blue and purple, respectively, and the conserved GXXG loop in each KH domain is boxed. Again, the red asterisk represents the location of Ile304. The secondary structure elements on top of the sequence alignment are placed with respect to the crystal structure of hFMRP(KH1-KH2Δ) and are labeled according to standard KH nomenclature, shown in Figure 2 and discussed in the main text. The black line connecting β2 and β′ in the second KH domain denotes the variable loop, and the amino acids absent in hFMRP(KH1-KH2Δ) are denoted by the dotted line.

more than 20 high-resolution structures of isolated KH domains have been reported (Berman et al., 2005), the only protein for which a structure of tandem KH domains has been published is the prokaryotic protein NusA, in which there are two tandem Type II KH domains (Gopal et al., 2001; Beuth et al., 2005).

Here, to our knowledge, we present the first structure of tandem, Type I KH domains from the eukaryotic protein FMRP (Figure 3A). The structure of the tandem KH domains of FMRP reveals that residue Ile304 is solvent inaccessible and is part of the hydrophobic core of the protein. Ile304 is therefore unlikely to be involved in direct contacts

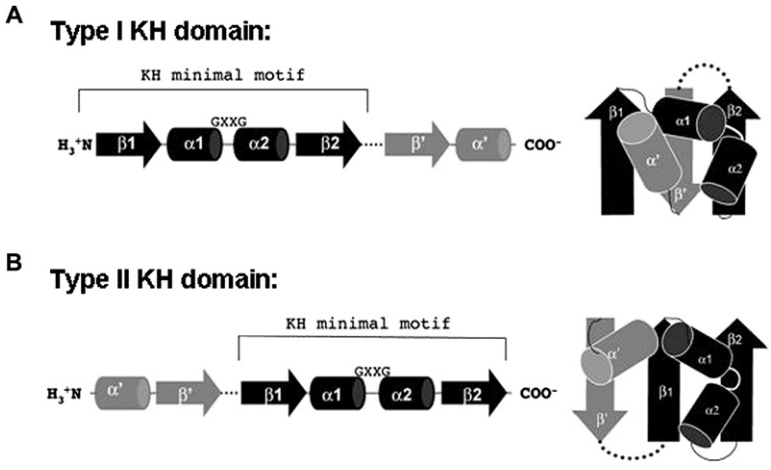

Figure 2. KH Fold

(A and B) Stylized representation of (A) the eukaryotic (Type I) KH domain and (B) the prokaryotic (Type II) KH domain. The labeling of secondary structure elements is done according to standard KH nomenclature. The dotted line connecting β2 and β′ represents the variable loop. The white line connecting helices α1 and α2 represents the GXXG loop.

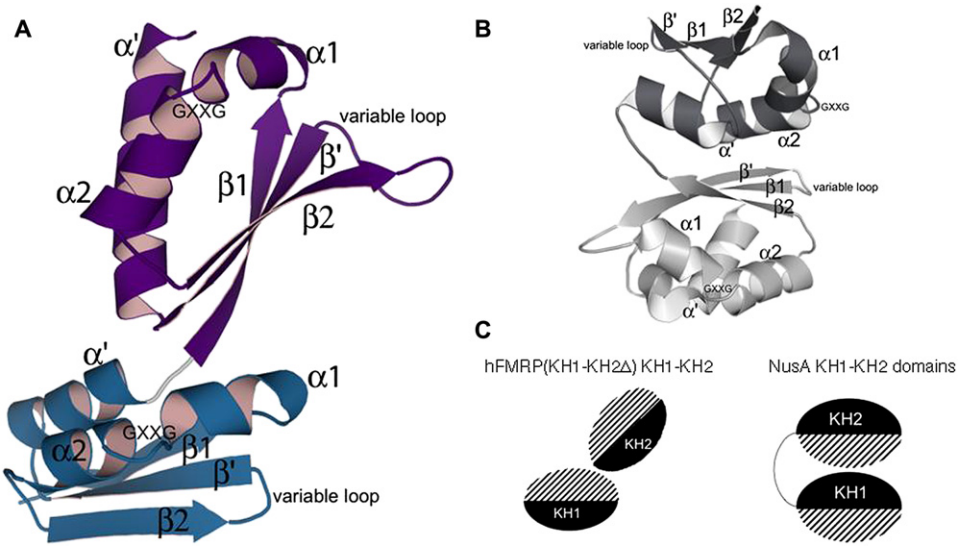

Figure 3. Crystal Structure of hFMRP(KH1-KH2Δ) and NusA(KH1-KH2)

(A) The KH1 and KH2 domains of hFMRP(KH1-KH2Δ) are colored blue and purple, respectively. See Table 1 for salient crystallographic statistics.

(B) NusA(KH1-KH2); PDB ID 1K0R. The KH1 and KH2 domains are colored light and dark gray, respectively.

(C) Cartoon representation of hFMRP(KH1-KH2Δ) and NusA(KH1-KH2) protein (PDB ID: 1K0R). Each KH domain is represented as an oval; the three-stranded β sheet is shaded in solid black, and the three α helices that pack against the β sheet are shaded with stripes.

to ligand, unless there are substantial structural rearrangements upon ligand binding. Furthermore, we show that mutation of Ile304 to Asn both perturbs the structure of the protein and decreases its stability.

RESULTS AND DISCUSSION

Structure of hFMRP(KH1-KH2Δ)

The crystallized construct has a shortened variable loop between β sheets β2 and β′ in KH2, and we refer to it as hFMRP(KH1-KH2Δ). We made many different constructs of the KH1-KH2 domains of human FMRP, FXRP1, FXRP2, and dFXRP, but only hFMRP(KH1-KH2Δ) gave diffraction-quality crystals. The sequence of this construct as well as its relationship to the *Drosophila* ortholog (dFXRP) and the human paralogs FXRP1 and FXRP2 are shown in Figure 1B. The structure was phased with multi-wavelength anomalous diffraction data from selenomethionine-substituted protein and was refined to R_{work} = 23.0 and R_{free} = 28.6, at 1.9 Å resolution (Table 1). The KH1 and KH2 domains of hFMRP(KH1-KH2Δ) both adopt a Type I fold, in which a β sheet composed of three antiparallel strands is abutted by three α helices (α1, α2, and α′). The main hydrophobic core of the domain comprises the buried hydrophobic residues between the hydrophobic faces of the β sheet and the α helices (Figure 3A). The β sheet in each KH domain, described in reference to standard KH nomenclature (Grishin, 2001), consists of two parallel, core β strands, β1 and β2, that sandwich the β′ strand (Figure 2A). This all-antiparallel arrangement of strands distinguishes the eukaryotic Type I KH fold from the prokaryotic Type II KH fold, in which β1 and β2 are adjacent and parallel to each other, but in which the β′ strand is adjacent and antiparallel to β1 (Figure 2B). The length and sequence of the variable loop are different in different KH domains. Although the variable loop in the KH2 domain of FMRP is truncated to ten amino acids in the construct crystallized, it still dramatically protrudes from the side of the domain, whereas the variable loop in KH1 is markedly shorter (Figure 3A).

Comparison of hFMRP(KH1-KH2Δ) and NusA KH Domains

Not only is the order of secondary structural units in the individual KH units of FMRP different from that of prokaryotic KH domains, but the relative orientation of the KH1 and KH2 domains in hFMRP(KH1-KH2Δ) is also dramatically different from the orientation of the tandem KH domains of bacterial protein NusA (Gopal et al., 2001; Beuth et al., 2005) (Figures 3B and 3C) (PDB ID: 1K0R). In NusA, an unstructured six amino acid linker connects KH1 to KH2, and an area of ~1380 $Å^2$ is buried at the interface between the β sheet of KH1 and the α helices (α′ and α2) of KH2. By contrast, in hFMRP(KH1-KH2Δ), the α′ helix of KH1 is linked to the β1 strand of KH2 by the single residue Glu280, which adopts non-β, non-α ϕ/ψ angles to accomplish this tight connection, and there are minimal contacts between the KH1 and KH2 domains (Figures 3A and 3C). These features are consistent with the fact that the KH1 and KH2 domains of FMRP and its relatives can be expressed separately.

The tandem arrangement of KH domains in FMRP has implications for protein function and ligand recognition. The structure of bacterial NusA in complex with its ligand shows an extended RNA, which makes contacts with residues in both the Type II KH1 and KH2 domains (Beuth

Table 1. Data Collection and Refinement

	Peak[a]	Inflection	Remote
Wavelength	0.9788	0.9793	0.9184
Completeness (%)	98.0 (83.8)	98.2 (85.9)	99.2 (95.4)
Resolution (Å)	50–1.9 (1.97–1.9)	50–1.9 (1.97–1.9)	50–1.9 (1.97–1.90)
Signal/noise (I/σ)	19.6 (4.8)	27.3 (3.7)	24.5 (2.9)
R_{sym}[b] (%)	5.0 (18.3)	3.4 (22.3)	3.5 (28.6)
Redundancy	4.6 (4.2)	4.7 (4.1)	4.8 (4.2)
Unique reflections	24,901	24,950	25,337
Refinement			
Resolution	1.9 Å		
Number of atoms	2,387		
Protein	2,246		
Waters	136		
Se	4		
Mg	1		
Rmsd			
Bond lengths (Å)	0.01		
Bond angles (°)	1.4		
Average B value	40.0		
$R_{working}$ (%)	23.0		
R_{free}[c] (%)	28.6		
Ramachandran plot (two complexes per asymmetric unit)			
Residues in the most favored regions	235		
Residues in the additionally allowed regions	15		
Residues in the generously allowed regions	0		
Residues in the disallowed regions	0		

The values in parentheses are for the highest-resolution bins.
[a] The model was refined against the peak data set.
[b] $R_{sym} = \Sigma_h\Sigma_i|I_{h,i} - \langle I\rangle_h|/|\Sigma_h\Sigma_i I_{h,i}|$, where $\langle I\rangle_h$ is the average intensity of symmetry-related reflections.
[c] R_{free} was calculated the same as R_{work}, but on 5% of randomly selected data not used in refinement.

et al., 2005). The structures of NusA with and without RNA bound show no significant changes in protein conformation upon ligand binding (Beuth et al., 2005; Gopal et al., 2001). The orientation of the tandem KH domains of FMRP is different from that seen in NusA; therefore, we would expect the mode of interaction with ligand to also be different.

Although no high-resolution structural data are available, it has been proposed that the *Drosophila* protein PSI interacts with its RNA substrate through an elongated interaction surface that extends across its four Type I KH domains (Chmiel, et al. 2006). In such a model, tandem KH domains function cooperatively. This observation suggests, again, that the orientation of KH domains with respect to each other is important in ligand binding (Figure 3C).

hFMRP(KH1-KH2Δ) Is a Monomer in Solution

It has been suggested that KH motifs may be able to homodimerize (Lewis et al., 1999, 2000; Git and Standart, 2002; Ramos et al., 2002), and that full-length FMRP may homodimerize through its N-terminal and KH2 domains (Adinolfi et al., 2003; Ramos et al., 2006), although the evidence is by no means conclusive. In the crystal structure we present, there are two molecules in the asymmetric unit related by noncrystallographic symmetry (NCS) that are close to perpendicular to the two-fold crystallographic B axis. The interface between the two molecules has a total buried surface area of 2137 $Å^2$ and is primarily formed between NCS-related β2 strands (residues 314–322) of the KH2 domain. This amount of buried surface area is well above the accepted minimum of 1200 $Å^2$ for a protein-protein dimerization interface (Lo Conte et al.,

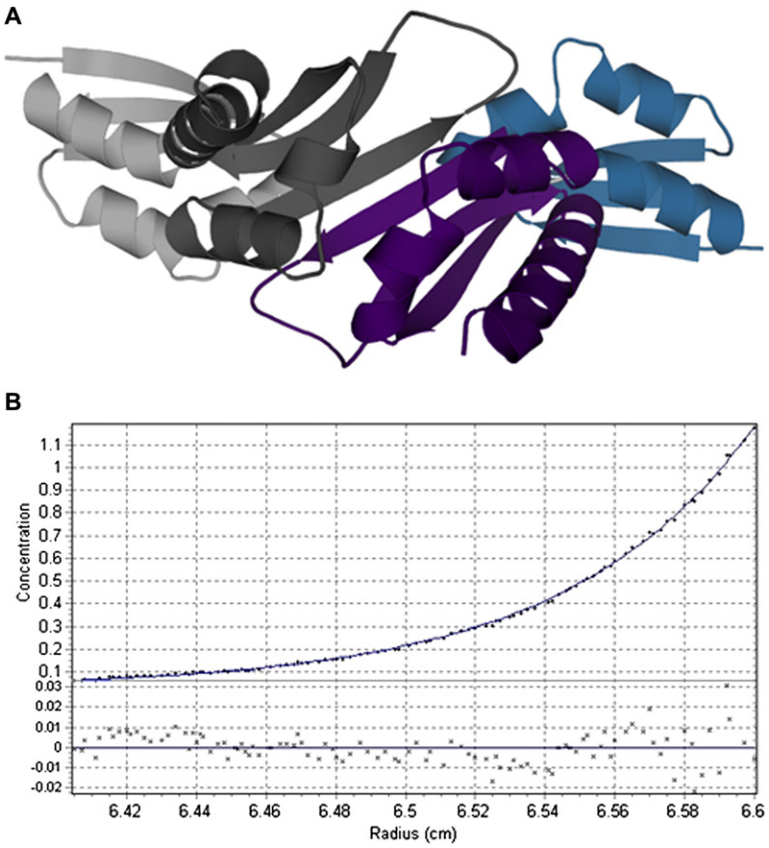

Figure 4. hFMRP(KH1-KH2Δ) Behavior in Solution

(A) The contents of the asymmetric unit (ASU). There are two molecules in the ASU related by NCS that are close to perpendicular to the crystallographic B axis. The KH1 and KH2 domains of one molecule are colored blue and purple, respectively. The NCS-related chain is colored in shades of gray.

(B) Representative example of analytical ultracentrifugation (AUC) data. Protein concentration is plotted versus radius for an AUC equilibrium experiment. 150 μM and 260 μM hFMRP(KH1-KH2Δ) were spun at 10,000, 18,000, 38,000, and 45,000 rpm at 25°C. The solid line shows a fit of the data to a model of a nonassociating monomer with a molecular weight of 15,967 (rmsd of 0.007). The molecular weight of the monomer calculated from its sequence is 16,037. Residuals are shown at the bottom of the plot.

1999). Moreover, the arrangement creates an extended β sheet plane composed of six antiparallel strands in the asymmetric unit. Peptide backbone hydrogen bonding and cross-strand side chain interactions stabilize the β sheet (Merkel, et al., 1999) (Figure 4A). Gel-filtration chromatography, however, showed no indication of dimerization. We were therefore motivated to determine the oligomeric state of hFMRP(KH1-KH2Δ) in solution by using more sensitive means and performed sedimentation equilibrium ultracentrifugation measurements over a range of protein concentrations and rotor speeds. These studies clearly showed a single species in solution with a molecular weight consistent with that predicted for the monomer (Figure 4B). Even at high protein concentrations there was no indication of higher-molecular weight species. Thus, the NCS dimer seen in the crystal structure of the KH1-KH2 domain construct does not represent a dimer that is stable in solution. This observation does not, of course, preclude the possibility that the full-length hFMRP is a dimer, or that there could be KH domain-mediated contacts in such a dimer.

Structural and Solution Characterization of Ile304Asn

There is special interest in residue Ile304, which is located in the KH2 domain of FMRP, because an individual with a particularly severe case of Fragile X syndrome produces normal levels of a mutant form of FMRP, with the Ile304Asn mutation (De Boulle et al., 1993). Since the original clinical description, there have been several hypotheses regarding the possible effect of the Ile304Asn mutation on protein function. These include the proposal that the Ile304Asn mutation causes complete unfolding of the KH2 domain (Musco et al., 1997); that the mutation has no effect on protein structure, but that it prevents RNA binding by disrupting a hydrophobic platform on the protein involved in RNA contact (Lewis et al., 2000); and, finally, that the mutation causes an alteration in the association of FMRP with other proteins in vivo (Feng et al., 1997). The controversy surrounding the effect of the mutation derives in part from the fact that data obtained with different KH domains have been extrapolated to the KH domains of hFMRP (Lewis et al., 2000; Pozdnyakova and Regan, 2005; Chmiel et al., 2006).

The structure of hFMRP(KH1-KH2Δ) that we present allows us to precisely specify the position of the Ile304 residue in the protein: Ile304 forms part of an extensive network of hydrophobic residues (Val296, Ile307, Val308, Val316, Ile318, Phe380, Phe382, and Leu395) that stabilize α1, α2, and α′ on the β sheet of the second KH domain (Figure 5A). This set of hydrophobic residues is conserved among all known Fragile X-related proteins, including FXRP1, FXRP2, and dFXRP, suggesting a conservation of hydrophobic packing and van der Waals interactions in the core. All of the atoms of Ile304 are completely inaccessible to solvent, except for the atom Ile304Cγ2, whose

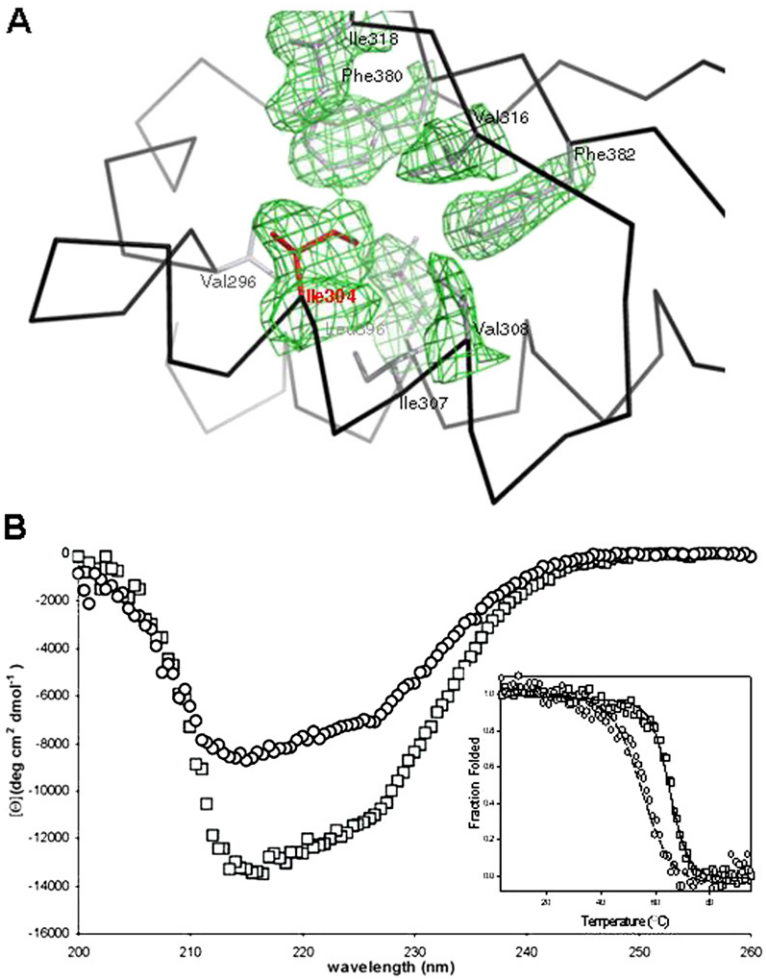

Figure 5. Structural Analysis of Ile304

(A) Stick representation of the hydrophobic network of amino acids with experimental electron density (contoured at $\sigma = 1.5$) after solvent flattening and NCS averaging with RESOLVE (Terwilliger, 2004). The electron density around residues Val296 and Ile307 is omitted for clarity. Ile304 is colored red.

(B) CD spectra of hFMRP(KH1-KH2Δ) (squares) and hFMRP(KH1-KH2Δ) Ile304Asn (circles). Inset: thermal denaturation curves of hFMRP (KH1-KH2Δ) (squares) and hFMRP(KH1-KH2Δ) Ile304Asn (circles). The solid and dotted lines show fits to a two-state denaturation transition for hFMRP(KH1-KH2Δ) and hFMRP (KH1-KH2Δ) Ile304Asn, respectively. We estimate the melting temperature of hFMRP(KH1-KH2Δ) as 65°C and that of hFMRP(KH1-KH2Δ) Ile304Asn as 55°C.

solvent accessibility is calculated to be less than a third of that of the Cγ2 atom of Ile in an extended Gly-Ile-Gly environment (CCP4, 1994). This observation indicates that Ile304 cannot participate in direct contacts with ligand, unless there are significant structural perturbations. Because Ile304 is an integral component of a buried network of hydrophobic amino acids, its substitution with Asn would disrupt the hydrophobic core and destabilize the protein.

We used circular dichroism (CD) to investigate the effect of the Ile304Asn mutation on the structure and stability of hFMRP(KH1-KH2Δ) in solution. The mutation causes both a decrease in the secondary structure content of the protein and a decrease in its stability (Figure 5B). The CD spectra of hFMRP(KH1-KH2Δ) Ile304Asn at 4°C and at 25°C are identical, indicating that the mutation does not shift the equilibrium between the folded and unfolded states of the protein, but rather that it causes a structural change in the protein (data not shown).

In summary, we present the structure of tandem, eukaryotic, Type I KH domains. The relative orientation of these tandem domains is quite different from that observed for tandem, prokaryotic, Type II domains, as epitomized in the structure of NusA. Most importantly, the structure of the tandem KH domains of FMRP has immediate relevance to our understanding of the molecular basis of Fragile X syndrome. Our results reveal the location of amino acid Ile304, whose mutation to Asn is associated with a particularly extreme case of Fragile X syndrome. Moreover, we show that mutation of Ile304 to Asn causes significant structural perturbation and destabilization of the protein in vitro, thus providing one plausible mechanism by which this mutation severely compromises protein function. Future studies will focus on a better understanding of the normal function of FMRP and the nature of its RNA ligands.

EXPERIMENTAL PROCEDURES

Molecular Biology

The hFMRP(KH-KH2Δ) construct corresponds to amino acids 216–404 of human FMRP (NCBI accession number: AAH86957), with

residues 331–375 deleted from the variable loop of the KH2 domain. This coding sequence was cloned into the modified vector pET15b (Novagen; Madison, WI) in which we incorporated an N-terminal His tag followed by a TEV protease site. The Ile304Asn mutant was generated from the hFMRP(KH-KH2Δ) construct, by using Quick Change mutagenesis (Stratagene; La Jolla, CA), and was verified by DNA sequencing.

Expression and Purification

Metal affinity purification for hFMRP(KH-KH2Δ) and hFMRP (KH-KH2Δ) Ile304Asn was adapted from previously published methods (Pozdnyakova and Regan, 2005). Briefly, plasmids were transformed into BL21 GOLD (DE3) and were grown with shaking in Luria-Bertani (LB) supplemented with 100 μg/ml ampicillin at 37°C to an OD_{600} of 0.6, at which point the expression was induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1mM. The temperature was then lowered to 25°C, and growth continued for 5 more hr.

Cells were harvested by centrifugation, resuspended in 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 5 mM β-mercaptoethanol, 10 mM Imidazole, 20% glycerol containing lysozyme (1 mg/ml) and complete EDTA-free protease inhibitor cocktail (Roche; Basel, Switzerland), and lysed by sonication. Debris was removed by centrifugation. The soluble supernatant fraction of the whole-cell lysate was incubated with TALON metal affinity resin (BD Biosciences Clonetech; Palo Alto, CA) for 45 min at 4°C. His-tagged protein was eluted with elution buffer (50 mM Tris-HCl [pH 8.0], 300 mM NaCl, 5 mM β-mercaptoethanol, 250 mM imidazole, 5% glycerol). The N-terminal His tag was then removed by digestion with TEV protease (Invitrogen; Carlsbad, CA). After TEV protease digestion, the sample was passed over a second TALON metal affinity column to remove the cleaved His tags and the TEV protease (which is also His tagged). Fractions of hFMRP(KH-KH2Δ) or hFMRP(KH-KH2Δ)-Ile304Asn with no tag were then pooled, concentrated (Centriprep YM-10; Fisher Scientific; Pittsburg, PA), and loaded onto a gel-filtration column (High-load Superdex RH-75 column, Amersham Biosciences; Piscataway, NJ) equilibrated with 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 5 mM L-glutathione. The identity and purity of the KH-containing fractions were checked by SDS-PAGE, and the fractions containing isolated protein sample were pooled and further concentrated by Centriprep-YM-10 to a final concentration of approximately 32–36 mg/ml. The protein concentration was determined from absorption at 280 nm (the extinction coefficient was determined by amino acid composition analysis by using the PROTPARAM tool; http://ca.expasy.org/ [Gasteiger et al., 2003]).

Selenomethionine (SeMet)-substituted hFMRP(KH1-KH2Δ) was produced as follows. The B834(DE3) *Escherichia coli* auxotroph was transformed with plasmid and grown in M9 media supplemented with 50 mg ml^{-1} L-methionine at 37°C. At an OD_{600} of 0.6 cells were spun down and resuspended in M9 media supplemented with 50 mg ml^{-1} of L-selenomethionine. IPTG was then added to a final concentration of 0.7 mM, and growth continued for an additional 17 hr at 30°C. Cell harvesting and protein purification proceeded as described for unlabeled protein.

Crystallization of hFMRP(KH1-KH2Δ)

Crystals of hFMRP(KH-KH2Δ) (wild-type and SeMet-substituted) grew rapidly (in ~8 hr) when the hanging-drop vapor-diffusion method was used. Protein concentration was 32–36 mg/ml in 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, and 5 mM reduced L-glutathione. The well solution consisted of 0.1 M Tris-HCl (pH 7.5–8.5), 0.2 M MgCl, and PEG 4000 27%–30%. Single crystals were prepared for data collection by fast transfer into cryoprotectant, Paratone-N (Hampton Research), and were immediately frozen in the cryostream. Crystals grew in the space group C2 with unit cell dimensions of a = 71.0 Å, b = 70.7 Å, c = 68.2 Å, and β = 107.2 Å; there were two molecules in the asymmetric unit.

Data Collection, Structure Determination, and Refinement

All data were collected at beamline X6A, National Synchrotron Light Source, Brookhaven from a single SeMet-substituted protein crystal (Table 1). Peak, remote, and inflection data sets were scaled and integrated with the HKL2000 program suite (Otwinowski and Minor, 1997). The structure was phased with SeMet multiwavelength anomalous diffraction (MAD) data at 2.5 Å by using SOLVE (Terwilliger, 2004). Four Se sites in the asymmetric unit were found with a figure of merit of 0.65. Solvent flattening, two-fold noncrystallographic symmetry averaging, and phase extension to 1.9 Å in RESOLVE (Terwilliger, 2004) produced an interpretable map that was used in building the structure. The model was traced in O (Jones et al., 1991) by using simulated annealing composite-omit maps calculated by using CNS (Brunger et al., 1998).

Water molecules were added in two cycles of ARP/wARP (CCP4, 1994) and were validated by using Fourier difference maps and COOT (Emsley and Cowtan, 2004). Iterative rounds of restrained refinement in REFMAC5 (CCP4, 1994) and model adjusting in COOT were carried out until R factors dropped to acceptable values for a structure of this resolution (Kleywegt and Jones, 2002). A TLS model (Painter and Merritt, 2006) was used in the late stages of refinement. The stereochemical quality of the model was inspected by using ProCheck and Molprobity (Lovell et al., 2003). Figures were made with the help of PyMOL (DeLano, 2002). Solvent-accessible areas were calculated with AREAIMOL (CCP4, 1994) (comparison with other structures was performed using DALI, see the Supplemental Data available with this article online).

Circular Dichroism

This protocol was adapted from a previously published method (Pozdnyakova and Regan, 2005). Briefly, CD spectra were recorded in a 0.1 cm path-length cuvette by using AVIV spectrophotometer Model 215 (AVIV Instruments, Inc.) at 25°C. Protein samples were diluted in CD buffer (10 mM HEPES [pH 7.5], 30 mM NaCl) to a final concentration of 25 μM, as determined by amino acid analysis. For each sample spectrum recorded, a buffer blank was subtracted from the raw signal, and, subsequently, mean residue ellipticity was calculated. Thermal denaturation transitions were monitored by CD absorption at 222 nm. Thermal scans were performed in forward and reverse directions from 4°C to 95°C in 1°C steps with a 3 min equilibration time at each temperature.

Analytical Ultracentrifugation

Sedimentation equilibrium was performed on a Beckman XL-I analytical ultracentrifuge at 25°C by using an AN 60-Ti 4-hole rotor equipped with six-channel, carbon-epoxy composite centerpieces (Beckman Coulter). hFRMP(KH1-KH2Δ) was resuspended in AUC buffer (150 mM NaCl, 50 mM HEPES [pH 8.0], 1 mM TCEP [pH 7.0] [Pierce; Rockford, IL]) at 150 μM and 260 μM. Approximately 120 μl of each sample was spun at 10,000, 18,000, 38,000, and 45,000 rpm and was allowed to reach equilibrium, which took about ~24 hr in each case. Sedimentation equilibrium curves were measured by absorbance at 280 nm, and successive scans were taken at 2 hr intervals. Attainment of equilibrium for every speed was confirmed by comparing the radial concentration profile in eight successive scans by using the MATCH program. For each sample, the various data sets were fitted both individually for each speed and simultaneously at all speeds by using HeteroAnalysis software (Cole, 2004) (V1.1.19, by James Cole and Jeffery Lary, Analytical Ultracentrifugation Facility, Biotechnology and Bioservices Center, University of Connecticut, Storrs, CT). Two fitting models were used; one is the ideal model, which assumes that the solution is ideal with single species, while the second model assumed monomer-dimer equilibria and a fixed monomer molecular weight of 16,037, as calculated from the amino acid sequences. The viscosity of the buffer at 25°C was calculated to be 1 mg/ml, and the partial specific volume for hFMRP(KH1-KH2Δ) was calculated on the basis of amino acid composition as 0.727 ml/mg.

Supplemental Data

Supplemental Data including two tables may be found with this article online at http://www.structure.org/cgi/content/full/15/9/1090/DC1/.

ACKNOWLEDGMENTS

We thank the staff at Brookhaven National Laboratory, National Synchrotron Light Source beamline X6A for their skilled assistance. We gratefully acknowledge the advice and expertise of Yong Xiong, Jimin Wang, Andrea Berman, and Karin Reinisch. We thank the Regan lab for suggestions and comments on the manuscript, and Fang Yi for her invaluable help in performing the analytical ultracentrifugation experiments.

Received: May 18, 2007
Revised: June 11, 2007
Accepted: June 12, 2007
Published: September 11, 2007

REFERENCES

Adinolfi, S., Ramos, A., Martin, S.R., Dal Piaz, F., Pucci, P., Bardoni, B., Mandel, J.L., and Pastore, A. (2003). The N-terminus of the fragile X mental retardation protein contains a novel domain involved in dimerization and RNA binding. Biochemistry *42*, 10437–10444.

Antar, L.N., Dictenberg, J.B., Plociniak, M., Afroz, R., and Bassell, G.J. (2005). Localization of FMRP-associated mRNA granules and requirement of microtubules for activity-dependent trafficking in hippocampal neurons. Genes Brain Behav. *4*, 350–359.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2005). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

Beuth, B., Pennell, S., Arnvig, K.B., Martin, S.R., and Taylor, I.A. (2005). Structure of a *Mycobacterium tuberculosis* NusA-RNA complex. EMBO J. *24*, 3576–3587.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol. Crystallgr. *50*, 760–763.

Chmiel, N.H., Rio, D.C., and Doudna, J.A. (2006). Distinct contributions of KH domains to substrate binding affinity of *Drosophila* P-element somatic inhibitor protein. RNA *12*, 283–291.

Cole, J.L. (2004). Analysis of heterogeneous interactions. Methods Enzymol. *384*, 212–232.

De Boulle, K., Verkerk, A.J., Reyniers, E., Vits, L., Hendrickx, J., Van Roy, B., Van den Bos, F., de Graaff, E., Oostra, B.A., and Willems, P.J. (1993). A point mutation in the FMR-1 gene associated with fragile X mental retardation. Nat. Genet. *3*, 31–35.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System (www.pymol.org).

Devys, D., Lutz, Y., Rouyer, N., Bellocq, J.P., and Mandel, J.L. (1993). The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X permutation. Nat. Genet. *4*, 335–340.

Eberhart, D.E., Malter, H.E., Feng, Y., and Warren, S.T. (1996). The fragile X mental retardation protein is a ribonucleoprotein containing both nuclear localization and nuclear export signals. Hum. Mol. Genet. *5*, 1083–1091.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Feng, Y., Absher, D., Eberhart, D., Brown, V., Malter, H., and Warren, S. (1997). FMRP associates with polyribosomes as an mRNP, and the I304N mutation of severe fragile X syndrome abolishes this association. Mol. Cell *1*, 109–118.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. *31*, 3784–3788.

Git, A., and Standart, N. (2002). The KH domains of *Xenopus* Vg1RBP mediate RNA binding and self-association. RNA *8*, 1319–1333.

Gopal, B., Haire, L.F., Gamblin, S.J., Dodson, E.J., Lane, A.N., Papavinasasundaram, K.G., Colston, M.J., and Dodson, G. (2001). Crystal structure of the transcription elongation/anti-termination factor NusA from *Mycobacterium tuberculosis* at 1.7 Å resolution. J. Mol. Biol. *314*, 1087–1095.

Grishin, N.V. (2001). KH domain: one motif, two folds. Nucleic Acids Res. *29*, 638–643.

Hammond, L.S., Macias, M.M., Tarleton, J.C., and Shashidhar Pai, G. (1997). Fragile X syndrome and deletions in FMR1: new case and review of the literature. Am. J. Med. Genet. *72*, 430–444.

Jin, P., and Warren, S.T. (2000). Understanding the molecular basis of fragile X syndrome. Hum. Mol. Genet. *9*, 901–908.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Khandjian, E.W., Huot, M.E., Tremblay, S., Davidovic, L., Mazroui, R., and Bardoni, B. (2004). Biochemical evidence for the association of fragile X mental retardation protein with brain polyribosomal ribonucleoparticles. Proc. Natl. Acad. Sci. USA *101*, 13357–13362.

Kleywegt, G.J., and Jones, T.A. (2002). Homo crystallographicus-quo vadis? Structure *10*, 465–472.

Lewis, H.A., Chen, H., Edo, C., Buckanovich, R.J., Yang, Y.Y., Musunuru, K., Zhong, R., Darnell, R.B., and Burley, S.K. (1999). Crystal structures of Nova-1 and Nova-2 K-homology RNA-binding domains. Structure *7*, 191–203.

Lewis, H.A., Musunuru, K., Jensen, K.B., Edo, C., Chen, H., Darnell, R.B., and Burley, S.K. (2000). Sequence-specific RNA binding by a Nova KH domain: implications for paraneoplastic disease and the fragile X syndrome. Cell *100*, 323–332.

Lo Conte, L., Chothia, C., and Janin, J. (1999). The atomic structure of protein-protein recognition sites. J. Mol. Biol. *285*, 2177–2198.

Lovell, S.C., Davis, I.W., Arendall, W.B., 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by Cα geometry: ϕ,ψ and Cβ deviation. Proteins *50*, 437–450.

Merkel, J.S., Sturtevant, J.M., and Regan, L. (1999). Side chain interactions in parallel β sheets: the energetics of cross-strand pairings. Structure *7*, 1333–1341.

Musco, G., Kharrat, A., Stier, G., Fraternali, F., Gibson, T.J., Nilges, M., and Pastore, A. (1997). The solution structure of the first KH domain of FMR1, the protein responsible for the fragile X syndrome. Nat. Struct. Biol. *4*, 712–716.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Painter, J., and Merritt, E.A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr. D Biol. Crystallogr. *62*, 439–450.

Pozdnyakova, I., and Regan, L. (2005). New insights into Fragile X syndrome. Relating genotype to phenotype at the molecular level. FEBS J. *272*, 872–878.

Ramos, A., Hollingworth, D., Major, S.A., Adinolfi, S., Kelly, G., Muskett, F.W., and Pastore, A. (2002). Role of dimerization in KH/RNA complexes: the example of Nova KH3. Biochemistry *41*, 4193–4201.

Ramos, A., Hollingworth, D., Adinolfi, D., Castets, M., Kelly, G., Frenkiel, T., Bardoni, B., and Pastore, A. (2006). The structure of the N-terminal domain of the fragile X mental retardation protein: a platform for protein-protein interaction. Structure *14*, 21–31.

Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G. (1993). The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. Nucleic Acids Res. *21*, 1193–1198.

Siomi, H., Siomi, M.C., Nussbaum, R.L., and Dreyfuss, G. (1994). Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome. Cell *77*, 33–39.

Tamanini, F., Willemsen, R., van Unen, L., Bontekoe, C., Galjaard, H., Oostra, B.A., and Hoogeveen, A.T. (1997). Differential expression of FMR1, FXR1 and FXR2 proteins in human brain and testis. Hum. Mol. Genet. *6*, 1315–1322.

Terwilliger, T.C. (2004). SOLVE and RESOLVE: automated structure solution, density modification and model building. J. Synchrotron Radiat. *11*, 49–52.

Wan, L., Dockendorff, T.C., Jongens, T.A., and Dreyfuss, G. (2000). Characterization of dFMR1, a *Drosophila melanogaster* homolog of the fragile X mental retardation protein. Mol. Cell. Biol. *20*, 8536–8547.

Zhang, Y.Q., Bailey, A.M., Matthies, H.J., Renden, R.B., Smith, M.A., Speese, S.D., Rubin, G.M., and Broadie, K. (2001). *Drosophila* fragile X-related gene regulates the MAP1B homolog Futsch to control synaptic structure and function. Cell *107*, 591–603.

Accession Numbers

Coordinates for the crystal structure of the KH1-KH2 domains from human Fragile X Mental Retardation Protein have been assigned the PDB ID code 2QND.

